ԵՐԵՎԱՆԻ ՊԵՏԱԿԱՆ ՀԱՄԱԼՍԱՐԱՆԻ ԳԻՏԱԿԱՆ ՏԵՂԵԿԱԳԻՐ
УЧЕНЫЕ ЗАПИСКИ ЕРЕВАНСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА

Բնական գիտություններ 1,1999 Естественные науки

Математика

УДК 517.597

А.З. АРАКЕЛЯН

О ХВОСТЕ РАСПРЕДЕЛЕНИЯ РАЗМАХА ВЫБОРКИ

В настоящей работе изучается асимптотическое поведение хвоста функции распределения случайной величины

$$W_\nu = \sup_{1\le k\le \nu} \xi_k - \inf_{1\le k\le \nu} \xi_k$$

определяющего размах выборки $(\xi_1, \cdots, \xi_\nu)$ **случайного объема** $\nu \ge 2$. **Целочисленная случайная величина** ν **не зависит от** $\{\xi_n\}$. **Пусть** $1-F(t)+F(-t)$, **где** F **– функция распределения** $\{\xi_n\}$, **правильно меняется и** $M\nu < +\infty$, **где** M **– знак математического ожидания.**

Доказано, что $\lim\limits_{t\to+\infty} \dfrac{1-W(t)}{1-F(t)+F(-t)} = M_\nu$, **где** $W(t)$ **– функция распределения величины** W_ν.

§1. Введение. Пусть $\{\xi_n\}$ – последовательность независимых одинаково распределенных (НОР) случайных величин (СВ) с функцией распределения (ФР) F, ν – не зависящий от $\{\xi_n\}$ случайный индекс с распределением $\{c_k\}$.

В математической статистике важной характеристикой выборки $(\xi_1, \cdots, \xi_n)$ объема $n > 1$ является ее размах[1]:

$$W_n = \max_{1\le k\le n} \xi_k - \min_{1\le i\le n} \xi_i, \; n > 1.$$

Размах выборки случайного объема $\nu > 1$ определяем равенством

$$W_\nu = \sup_{1\le k\le \nu} \xi_k - \inf_{1\le i\le \nu} \xi_i .$$

Обозначим $W_n(t) = P(W_n < t), n > 1,\; W(t) = P(W_\nu < t), t \in R^+$, где P – знак вероятности.

Цель настоящей работы заключается в изучении асимптотического поведения хвоста $1-W(t)$ при $t \to +\infty$ в предположении

$$M_\nu = \sum_{k\ge 1} k \cdot c_k < +\infty$$

и правильного изменения (см.[2]) суммы хвостов $1-F(t)+F(-t)$ при $t \to +\infty$. Здесь M – знак математического ожидания.

Основной результат работы содержится в следующей теореме.

Теорема. Если $M_\nu < +\infty$ и $1 - F(t) + F(-t)$ при $t \to +\infty$ правильно меняется, то существует предел

$$\lim_{t \to +\infty} \frac{1 - W(t)}{1 - F(t) + F(-t)} = M_\nu .$$

Таким образом $1 - W(t)$ при $t \to +\infty$ правильно меняется.

Обозначим

$$S_n = \xi_1 + \cdots + \xi_n, n \geq 1, S_\nu = \xi_1 + \cdots \xi_\nu, R(t) = P(S_\nu < t), t \in R^1 .$$

При доказательстве теоремы используется ([3], гл. 8, зад.31).

Лемма 1. Если $M_\nu < +\infty$ и $1 - F(t)$ при $t \to +\infty$ правильно меняется, то существует предел

$$\lim_{t \to +\infty} \frac{1 - R(t)}{1 - F(t)} = M_\nu .$$

Вначале устанавливается частный случай $P(\nu = n) = 1$ теоремы.

Лемма 2. Если $1 - F(t) + F(-t)$ при $t \to +\infty$ правильно меняется, то

$$1 - W_n(t) \sim n\{1 - F(t) + F(-t)\}, t \to +\infty, n > 1^* . \tag{1}$$

Доказательство леммы 2 состоит из двух частей. В первой - дан прямой вывод (1) на базе известной формулы ([4], с.534-535):

$$W_n(t) = n \cdot \int_{-\infty}^{+\infty} \{F(x + t) - F(x)\}^{n-1} dF(x), \ t > 0, n > 1, \tag{2}$$

в предположении правильного изменения правого $1 - F(t)$ и левого $F(-t)$ хвостов при $t \to +\infty$. Из этого предположения следует правильное изменение $1 - F(t) + F(-t)$, однако обратное утверждение, вообще говоря, неверно. Вторая часть опирается на результаты первой и на двухсторонние оценки.

В изложении неоднократно обращение к следующей формуле Феллера [3], с. 319-320. Именно пусть $\{\xi_n\}$ - последовательность независимых СВ с ФР $\{F_n\}$ соответственно.

Если $1 - F_k(t)$ при $t \to +\infty$ и $k = \overline{1, n}$ правильно меняются, то

$$P(S_n \geq t) \sim \sum_{k=1}^{n} \{1 - F_k(t)\}, t \to +\infty, n \geq 1 . \tag{3}$$

Формула (3) установлена в [3] в частном случае, однако вероятностное доказательсто Феллера (как легко убедиться) проходит и в общем случае. Более того, из доказательства Феллера следует оценка

$$P(S_n \geq t) \gtrsim \sum_{k=1}^{n} \{1 - F_k(t)\}, t \to +\infty, n \geq 1^{**} , \tag{4}$$

без каких-либо ограничений на F_k, $k = \overline{1, n}$.

Мы предлагаем аналитический метод доказательства формулы (3), основанный на формуле свертки и ее факторизации.

§ 2. Формула Феллера. Вначале отметим, что формула Феллера содержит

---

* $f \sim g$ означает $(f / g) \to 1$.

** $f(t) \gtrsim g(t), t \to +\infty$ означает $\underline{\lim}_{t \to +\infty} (f(t) / g(t)) \geq 1$.

Следствие. Если $1-F_k(t)+F_k(-t)$ при $t\to+\infty$ и $k=\overline{1,n}$. правильно меняются, то

$$P(|S_n|\geq t)\sim\sum_{k=1}^{n}\{1-F_k(t)+F_k(-t)\},\, t\to+\infty.\, n\geq 1.$$

Действительно, F_k и ФР СВ $|\xi_k|$ при каждом $k=\overline{1,n}$ имеют одинаковую сумму хвостов. Поэтому, по формуле (3), при $t\to+\infty$

$$P(|S_n|\geq t)\leq P\left(\sum_{k=1}^{n}|\xi_k|\geq t\right)\sim\sum_{k=1}^{n}P(|\xi_k|\geq t)=\sum_{k=1}^{n}\{1-F_k(t)+F_k(-t)\}. \quad (5)$$

Так как $F_k(-t), k=\overline{1,n}$ – правый хвост ФР СВ $(-\xi_k)$, то без ограничений на F_k (см.(4)) справедлива оценка

$$P(S_n\leq -t)=P(-S_n\geq t)\geq\sum_{k=1}^{n}P(-\xi_k\geq t)=\sum_{k=1}^{n}F_k(-t),\ t\to+\infty,$$

откуда и из (4) получаем

$$P(|S_n|\geq t)\geq\sum_{k=1}^{n}\{1-F_k(t)+F_k(-t)\},\ t\to+\infty. \quad (6)$$

Следствие вытекает из (5) и (6). ▷

Докажем (3), предположив в начале непрерывность F_k в нуле при каждом $k\geq 1$.

При $0<\rho_i<1$, где $\rho_i=P(\xi_i\leq 0)$, $i=1,2$, строим ФР:

$$F_i^-(x)=\begin{cases}1, & \text{если } x\geq 0,\\ \rho_i^{-1}\cdot F_i(x), & \text{если } x<0.\end{cases}\quad \text{и}\quad F_i^+(x)=\begin{cases}\dfrac{F_i(x)-\rho_i}{1-\rho_i}, & \text{если } x\geq 0,\\ 0, & \text{если } x<0,\end{cases} \quad (7)$$

для некоторых СВ $\xi_i^-\leq 0$ и $\xi_i^+\geq 0$, где векторы (ξ_1^-,ξ_1^+) и (ξ_2^-,ξ_2^+) независимы.

Тогда $F_i(x)=\rho_i F_i^-(x)+(1-\rho_i)F_i^+(x), x\in R^1$.

Если $\rho_i=1(\rho_i=0)$, то полагаем $F_i=F_i^-(F_i=F_i^+)$.

Из формулы свертки вытекает представление

$$P(S_2\geq x)=\rho_i\cdot(1-\rho_2)\{1-F_1^-*F_2^+(x)\}+(1-\rho_1)\cdot\rho_2\{1-F_1^+*F_2^-(x)\}+$$
$$+(1-\rho_1)(1-\rho_2)\{1-F_1^+*F_2^+(x)\},\, x>0. \quad (8)$$

Здесь $*$ – знак свертки.

Случай 1. $\xi_1\geq 0, \xi_2\geq 0$. Из формулы свертки следует неравенство

$$P(S_2\geq x)=1-F_2(x)+\int_0^x\{1-F_1(x-y)\}dF_2(y)\geq$$
$$\geq 1-F_2(x)+\{1-F_1(x)\}\cdot F_2(x),\, x>0,$$

откуда выводим (4) при $n=2$. Сама формула свертки при любом $\delta\in(0,1)$ представима в виде

$$P\left(S_2 \geq x\right) = 1 - F_2(x) + A_\delta(x) + B_\delta(x),\, x = 0, \tag{9}$$

где
$$A_\delta(x) = \int_0^{1-\delta} \left\{1 - F_1\left(x \cdot (1-u)\right)\right\} dF_2(xu),$$

$$0 \leq B_\delta(x) = \int_{1-\delta}^{1} \left\{1 - F_1\left(x \cdot (1-u)\right)\right\} dF_2(xu) \leq F_2(x) - F_2\left((1-\delta)x\right).$$

Если $1 - F_1(t)$ при $t \to +\infty$ правильно меняется, то равномерно по $v \in [a, b],\, 0 < a < b < +\infty$, существует предел [2]

$$\lim_{t \to +\infty} \frac{1 - F_1(v \cdot t)}{1 - F_1(t)} = v^{-\alpha_1}$$

при некотором $\alpha_1 \geq 0$. Следовательно, в силу предельного соотношения

$$\lim_{x \to +\infty} \int_0^{1-\delta} (1-u)^{-\alpha_1} dF_2(xu) = 1$$

для любого $\zeta \in (0,1)$ найдется $x_0 > 0$ такое, что

$$\left| \frac{A_\delta(x)}{1 - F_1(x)} - 1 \right| < \zeta \text{ при всех } x \geq x_0.$$

В силу (9) и оценок для A_δ и B_δ при $x \geq x_0$ получаем

$$\left\{1 - F_1(x)\right\} \cdot (1 - \zeta) + \left\{1 - F_2(x)\right\} \leq P\left(S_2 \geq x\right) \leq$$
$$\leq \left\{1 - F_1(x)\right\}(1 + \zeta) + \left\{1 - F_2\left((1-\zeta)\right)x\right\},$$

откуда следует (3) при $n = 2$.

Случай 2. $\xi_1 \geq 0, \xi_2 \leq 0$. Этот случай изучается аналитически аналогично предыдущему. Однако проще использовать вероятностные соображения.

Метод Феллера дает (4) при $n = 2$. Из-за $\xi_2 \leq 0$ имеем $P\left(S_2 \geq t\right) \leq P\left(\xi_1 \geq t\right)$ при любом $t > 0$, откуда и из (4) следует (3) при $n = 2$.

Теперь формула (3) в случаях 1 и 2, подставленная в (8), приводит к формуле

$$P\left(S_2 \geq x\right) \sim \rho_1\left(1 - \rho_2\right) \cdot \left\{1 - F_2^+(x)\right\} + \left(1 - \rho_1\right) \cdot \rho_2 \left\{1 - F_1^+(x)\right\} +$$
$$+ \left(1 - \rho_1\right)\left(1 - \rho_2\right) \sum_{i=1}^{2} \left\{1 - F_i^+(x)\right\},\; x \to +\infty.$$

Подстановка сюда вида F_i^+ из (7) дает (3) при $n = 2$. Индукцией утверждение распространяется на $n > 2$. ⊳

Распространим результат на случай, когда F_i имеют скачки в нуле.

Пусь F_1 имеет скачок в нуле, т.е. $F_1(+0) - F_1(0) = p_1 > 0$, а F_2 непрерывна в нуле. Обозначим

$$E(x) = \begin{cases} 1, \text{если } x > 0, \\ 0, \text{если } x \leq 0, \end{cases} \quad M(x) = \begin{cases} \dfrac{F_1(x) - p_1}{1 - p_1}, \text{если } x > 0, \\ (1 - p_1)^{-1} \cdot F_1(x), \text{если } x \leq 0. \end{cases}$$

ФР M непрерывна в нуле.

Справедливо представление

$$F_1(x) = p_1 \cdot E(x) + (1 - p_1) \cdot M(x), x \in R^1.$$

По формуле свертки

$$P(S_2 \geq x) = p_1 \cdot \{1 - E * F_2(x)\} + (1 - p_1)\{1 - M * F_2(x)\}, x > 0.$$

Так как $$1 - M(x) = (1 - p_1)^{-1} \cdot \{1 - F_1(x)\}, x > 0,$$

$$p_1 \cdot \{1 - E * F_2(x)\} = p_1 \cdot \{1 - F_2(x)\},$$

То, по вышедоказанному,

$$P(S_2 \geq x) \sim p_1\{1 - F_2(x)\} + (1 - p_1)\{1 - M(x)\} + (1 - p_1)\{1 - F_2(x)\} =$$

$$\sum_{i=1}^{2}(1 - F_2(x)), x \to +\infty,$$

если $1 - F_1(x)$, либо $1 - F_2(x)$ при $x \to +\infty$ правильно меняется. Если $F_i, i = 1{,}2$ имеют скачки в нуле, то сохраняется приведенное доказательство с использованием предыдущего результата.

§ 3. Размах выборки. Лемма 2 при $n = 2$ может быть получена из следующих соображений. Размах $W_n, n > 1$, удовлетворяет случайному равенству

$$W_n = \max_{1 \leq i < j \leq n} |\xi_j - \xi_i|. \tag{10}$$

По следствию,

$$P(|\xi_2 - \xi_1| \geq t) \sim 2 \cdot \{1 - F(t) + F(-t)\}, t \to +\infty,$$

что в силу (10) доказывает (1) при $n = 2$. $\triangleright$.

Осуществление этой идеи в (10) затруднено при $n > 2$ из-за зависимости СВ $|\xi_j - \xi_i|, 1 \leq i < j \leq n$.

Поэтому мы обратимся к равенству (2). Смысл (2) раскрывает интегральная запиеь

$$W_n(t) = \int_{-\infty}^{+\infty} P\left(\max_{1 \leq k \leq n} \xi_k < x + t \,\middle|\, \min_{1 \leq i \leq n} \xi_i = x\right) dP\left(\min_{1 \leq i \leq n} \xi_i < x\right)$$

равенства

$$W_n(t) = P\left(\max_{1 \leq k \leq n} \xi_k - \min_{1 \leq i \leq n} \xi_i < t\right), t > 0, n > 1.$$

Здесь $P(A|B)$ – условная вероятность события A при условии осуществления события B.

Именно: из $\xi_1, \cdots, \xi_n$ выбирается n-способами одна, вероятность попадания которой в $[x, x + dx)$ есть $dF(x)$, а остальные помещаются в промежутке $[x; x + t)$ с вероятностью $\{F(t) - F(-t)\}^{n-1}$. $\triangleright$

Введем обозначения

$$D_{k,n}(t) = \int_{-\infty}^{+\infty} G_{n-k}(x + t) dG_k(x), k = \overline{1, n-1}, n > 1,$$

где $G_k(t) = P\left(\min\limits_{1\le i\le k} \xi_i < t\right)$. Нетрудно видеть, что

$$1 - D_{k,n}(t) = P\left(\min_{k<i\le n} \xi_i - \min_{1\le j\le k} \xi_j \ge t\right), t \in R^1, k = \overline{1, n-1}, n > 1. \quad (11)$$

Принимая во внимание равенства $n \cdot C_{n-1}^{k} = (k+1) \cdot C_n^{k+1}$ и

$$1 = n \cdot \int_{-\infty}^{+\infty} \{1 - F(x)\}^{n-1} dF(x) \left(= \int_{-\infty}^{+\infty} dG_n(x)\right),$$

преобразуем формулу (2):

$$1 - W_n(t) = n \cdot \int_{-\infty}^{+\infty} \left\{ (1 - F(x))^{n-1} - \left[(1 - F(x)) - (1 - F(x+t))\right]^{n-1} \right\} dF(x) =$$

$$n \cdot \sum_{k=0}^{n-2} C_{n-1}^{k} \cdot (-1)^{n-k} \int_{-\infty}^{+\infty} (1 - F(x))^{k} \cdot (1 - F(x+t))^{n-1-k} dF(x) =$$

$$\sum_{k=0}^{n-2} C_n^{k+1} \cdot (-1)^{n-k} \int_{-\infty}^{+\infty} \{1 - G_{n-k-1}(x+t)\} dG_{k+1}(x)$$

или

$$1 - W_n(t) = \sum_{k=1}^{n-1} C_n^{k} (-1)^{n-k-1} \{1 - D_{k,n}(t)\}, t > 0, n > 1. \quad (12)$$

Если $1 - F(t)$ и $F(-t)$ при $t \to +\infty$ правильно меняются, то для любых целых $k, 1 \le k \le n-1$, и $n > 1$ функции $(1 - F(t))^{n-k}$ и $k \cdot F(t)$ при $t \to +\infty$ правильно меняются. По формуле полной вероятности

$$G_k(-t) = F(-t) \cdot \sum_{i=0}^{k-1} (1 - F(-t))^{i} \sim k \cdot F(-t),\ t \to +\infty,\ k \ge 1,$$

и, значит, $G_k(-t)$ при $t \to +\infty$ правильно меняется.

Правые хвосты ФР СВ $\min\limits_{k<i\le n} \xi_i$ и $\left(-\min\limits_{1\le j\le k} \xi_j\right)$ равны

$$1 - G_{n-k}(t) = (1 - F(t))^{n-k} \text{ и } G_k(-t) \sim k \cdot F(-t), \quad t \to +\infty,$$

соответственно. Поэтому согласно (11), по формуле (3), имеем

$$1 - D_{k,n}(t) \sim 1 - G_{n-k}(t) + G_k(-t) \sim (1 - F(t))^{n-k} + k \cdot F(-t),$$

$$t \to +\infty,\ k = \overline{1, n-1}, n > 1.$$

С учетом тождеств

$$k \cdot C_n^{k} = n \cdot C_{n-1}^{k-1} \text{ и } \sum_{k=1}^{n-1} C_{n-1}^{k-1} \cdot (-1)^{n-k-1} = 1$$

при $t \to +\infty$ из (12) находим

$$1 - W_n(t) \sim \sum_{k=1}^{n-1} C_n^{k} \cdot (-1)^{n-k-1} \cdot \left\{ (1 - F(t))^{n-k} + k \cdot F(-t) \right\} \sim$$

$$\sim n \cdot (1 - F(t)) + \left\{ \sum_{k=1}^{n-1} k \cdot C_n^{k} \cdot (-1)^{n-k-1} \right\} \cdot F(-t) =$$

$$= n \cdot \{1 - F(t) + F(-t)\}, n > 1. \ \triangleright$$

Докажем (1) в предположении правильного изменения суммы хвостов $1-F(t)+F(-t)$ при $t \to +\infty$.

На основе выборки $(\xi_1, \cdots, \xi_n)$ объема $n>1$ с размахом W_n построим выборку $(|\xi_1|, \cdots, |\xi_n|)$ и обозначим через $\hat{W}_n$ ее размах. НОР СВ $|\xi|_k$, $k=\overline{1,n}$ сосредоточены на R^+, имеют ФР $F(t)-F(-t)$ и единственный правый хвост $1-F(t)+$ $+F(-t), t>0$. Поэтому для ФР $\hat{W}_n(t)=P(\hat{W}_n<t)$ справедлива формула (1):

$$1-\hat{W}_n(t) \sim n \cdot \{1-F(t)+F(-t)\},\ t \to +\infty, n>1.$$

С другой стороны, очевидно $\hat{W}_n \le W_n$ для любого $n>1$, что позволяет записать нижнюю оценку

$$1-W_n(t) \ge n \cdot \{1-F(t)+F(-t)\},\ t \to +\infty, n>1. \tag{13}$$

Верхняя оценка вытекает из неравенств

$$1-W_n(t) \le P\left(\max_{1 \le i<j \le n}\left(|\xi_i|+|\xi_j|\right) \ge t\right) \le P\left(\sum_{k=1}^{n}|\xi_k| \ge t\right) \sim$$
$$\sim n \cdot \{1-F(t)+F(-t)\},\ t \to +\infty \tag{14}$$

(использована формула (3)). Эта оценка вместе с (13) доказывает лемму 2. ▷

Докажем теорему. По формуле полной вероятности

$$1-W(t)+\sum_{n \ge 2} c_n \cdot (1-W_n(t)), t>0.$$

С учетом (14) запишем верхнюю оценку

$$1-W(t) \le \sum_{n \ge 2} c_n \cdot P\left(\sum_{k=1}^{n}|\xi_k| \ge t\right), t>0,$$

где из-за условия $\nu>1$ с вероятностью единица $(c_0=c_1=0)$ суммирование справа можно начать с $n=0$. Поэтому, по формуле полной вероятности,

$$1-W(t) \le P\left(\sum_{k=1}^{\nu}|\xi_k| \ge t\right), t>0.$$

Так как $\{|\xi_k|\}$ – НОР СВ с единственным (правым) хвостом $1-F(t)+F(-t)$, то, по лемме 1,

$$1-W(t) \le (1-F(t)+F(-t)) \cdot M\nu, t \to +\infty. \tag{15}$$

С другой стороны, по заранее данному $\zeta \in (0,1)$ выберем целое $m>1$ такое, что

$$M\nu-\sum_{n=1}^{m} n \cdot c_n < \zeta \cdot M\nu.$$

Принимая во внимание лемму 2 и это неравенство, выписываем нижнюю оценку

$$1-W(t) \ge \sum_{n=2}^{m} c_n \cdot (1-W_n(t)) \sim \{1-F(t)+F(-t)\} \cdot \sum_{n=1}^{m} n \cdot c_n >$$
$$>(1-\zeta) \cdot M\nu \cdot \{1-F(t)+F(-t)\}, t \to +\infty.$$

Устремляя $\zeta \downarrow 0$, получаем

$$1-W(t)\geq M\nu\cdot\{1-F(t)+F(-t)\},\ t\to+\infty.$$

Это неравенство и (15) доказывают теорему. ⊳

Приношу благодарность проф. Э.А. Даниеляну за постановку задачи и ценные указания.

Кафедра теории вероятностей
и математической статистики

Поступила 03.11.1998

ЛИТЕРАТУРА

1. **Дейвис Г.** Порядковые статистики, М., Наука, 1979, 335 с.
2. **Сенета Е.** Правильно меняющиеся функции, М., Наука, 1985, 141 с.
3. **Феллер В.** Введение в теорию вероятностей и ее приложения, М., Мир, 1984, т. 2, 751 с.
4. **Королюк В.С., Портенко Н.И., Скороход А.В., Турбин А.Ф.** Справочник по теории вероятностей и математической статистике, М., Наука, 1985, 640 с.

Ա.Զ. Առաքելյան

ԸՆՏՐԱՆՔԻ ԼԱՅՆՈՒԹՅԱՆ ԲԱՇԽՄԱՆ ՊՈՉԻ ՄԱՍԻՆ

Ա մ փ ո փ ու մ

Դիցուք $\{\xi_n\}$-ը անկախ, միևնույն բաշխված պատահական մեծությունների հաջորդականություն է F բաշխման ֆունկցիայով, ν-ն պատահական ինդեքս է և $\nu > 1$՝ մեկ հավանականությամբ:

Նշանակենք

$$W(t)=P\left(\sup_{1\leq k\leq \nu}\xi_k-\inf_{1\leq i\leq \nu}\xi_i<t\right),\ t>0,$$

որտեղ P-ն հավանականության նշանն է:

Աշխատանքի հիմնական արդյունքը հետևյալն է:

Եթե ν-ն կախված չէ $\{\xi_n\}$-ից, $M\nu$ մաթեմատիկական սպասումը վերջավոր է և $1-F(t)+F(-t)$-ն կանոնավոր է փոփոխվում, երբ $t\to+\infty$, ապա գոյություն ունի սահմանը.

$$\lim_{t\to+\infty}\frac{1-W(t)}{1-F(t)+F(-t)}=M\nu:$$